

Hanna Knäusl, Barbara Rösch und Lea Schubart, Regensburg

Einfluss von Kontextfaktoren auf das Suchverhalten in der Wikipedia

Artikel in der Wikipedia enthalten nicht nur Text, sondern auch Bilder, Grafiken, Listen, Infoboxen und vieles mehr. Um zu untersuchen, wie Nutzer diese Inhalte verwenden, fand im SS 2012 eine Datenerhebung statt. Ziel der Studie ist es, die Vorlieben der Nutzer in unterschiedlichen Suchszenarien herauszufinden und die empirische Forschung in diesem Bereich zu erweitern.

Deskriptoren: Benutzerforschung, empirische Untersuchung, Informationssuche, Bild, Text

Detecting context factors in Wikipedia search tasks

Articles in information portals like Wikipedia contain not only text, but also images, graphics, lists and navigational information. To investigate the usage of those content elements in different search scenarios we describe a lab study aiming at identifying the users' preferences for varying media types in the different search situations.

Keywords: user research, information search, image, text, empirical study

Influence des facteurs contextuels sur le comportement de recherche dans Wikipedia

Les articles dans Wikipédia contiennent non seulement du texte, mais aussi des images, des graphiques, des listes, des infoboxes et bien plus encore. Afin d'examiner comment les gens utilisent ces contenus, une collecte de données a lieu en été 2012. L'objectif de l'étude est de connaître les préférences des utilisateurs dans différents scénarios de recherche et de développer la recherche empirique dans ce domaine.

Mots-clés: étude des besoins des utilisateurs, recherche empirique, recherche d'information, image, texte

1 Einleitung

Das Informationsportal Wikipedia ist eine große Informationsquelle für viele unterschiedliche Bedürfnisse. Die breite Streuung der dort vertretenen Themengebiete und die technischen Freiheiten und Möglichkeiten bieten eine große inhaltliche und mediale Vielfalt. Die Artikel enthalten nicht nur Text, sondern auch Bilder, Grafiken, Listen, Infoboxen und vieles mehr. Um die Nutzung dieser Inhalte in unterschiedlichen Suchszenarien zu untersuchen, fand im SS 2012 im Rahmen des Seminars *Human Activity Recognition* eine Datenerhebung zur Klärung verschiedener Forschungsfragen statt. Die Fragestellungen stehen in drei Kontexten: Der erste Kontext ist die Bildverwendung, hier konkret im Zusammenhang mit Hilfe von implizitem Feedback zur adaptierten Inhaltsdarstellung, die den zweiten Kontext ausmacht. Um die Adaption von Website-Inhalten zu optimieren, geht es drittens um eine Untersuchung der Nutzerzufriedenheit in Abhängigkeit von bestimmten Inhaltselementtypen. Die Auswahl der verschiedenen Methoden zur Datenerhebung in diesem Experiment hat den Hintergrund, dass zwar Fragen zur Bildverwendung durchaus in Laborsituationen untersucht werden können, aber die Gewinnung von Daten zur kontextsensitiven Adaption von Inhalten einer Website ist eigentlich nur in Feldversuchen möglich. Methodisch war es in dieser Studie deshalb ein Ziel, durch die Kombination verschiedener Methoden Wege zu finden, wie ohne Laborausstattung, wie dem Eye-Tracker, trotzdem implizites Feedback zu den angebotenen Inhalten gewonnen werden kann.

2 Forschungskontext

2.1 Implicit Feedback

Im Rahmen des Information Retrieval wird auch die Gewinnung von implizitem Feedback durch die Auswertung der Interaktionen des Nutzers während der Suche in einem Informationssystem untersucht (z. B. White & Kelly, 2006). Scrollen (Claypool, 2001), Klick-Verhalten (Joachims, 2007) und die Augenbewegung während der Suche (Bu-

scher, 2008b) können wichtige Hinweise auf die aktuelle Zufriedenheit des Nutzers mit den Suchergebnissen liefern. Kann man also Nutzerverhalten bereits während der Interaktion mit dem System hinsichtlich der Zufriedenheit und der Relevanz der Inhalte analysieren, ist es vorstellbar, dass ein System dann auf Basis des gewonnenen Feedbacks individuell auf den Nutzer reagieren könnte.

2.2 Wissenserwerb mit Bildern

Im visuellen Zeitalter werden Informationen nicht mehr nur rein textlich beschrieben, sondern immer häufiger mit Hilfe von Bildern dargestellt. Besonders für den Wissenserwerb mit Medien (Wikipedia) wird Bildern eine große Bedeutung zugemessen. Die empirische Forschung zum Thema Wissenserwerb mit Text und Bild bestätigt die lernförderliche Wirkung von Illustrationen. Werden Bilder zu Texten hinzugefügt, zeigen sich bessere Verstehens- und Behaltensleistungen. Allerdings ist in Frage zu stellen, ob dies für jede Nutzersituation respektive in diesem Kontext für jeden Task-Typ (und jedes Suchziel) zutrifft.

2.3 Analyse von Nutzerverhalten und Nutzerzufriedenheit

Eine umfassende Analyse des Nutzerverhaltens soll es ermöglichen, anhand der erhobenen Interaktionsdaten den Task-Typ den der Nutzer ausführt, vorherzusagen. Für jeden, der im Experiment vorgegebenen Task-Typen (*Lookup*, *Learn* und *Casual-Leisure*) sind zur Lösung der Aufgabe andere Inhalte relevant. Die Daten des Firefox-Logs und des Eye-Trackers stehen im Mittelpunkt der Analyse.

Neben der Analyse des Nutzerverhaltens soll die Zufriedenheit des Nutzers in Abhängigkeit verschiedener Inhaltselemente untersucht werden. Angesichts der Vielzahl der Inhaltselemente in der Wikipedia erfolgt zunächst eine Beschränkung auf Bilder. Um eine Zufriedenheit des Nutzers zu erkennen, ist grundsätzlich die Analyse aller erhobenen Daten möglich. Im Moment werden aber vor allem die Daten des Elektromyographen und der Fragebögen verwendet, bzw. zunächst nur die Fragebögen herangezogen. Der Nutzer muss nach Ende jedes Tasks seine Zufriedenheit mit der Suche in eine Likert-Skala von „unzufrieden“ (1) bis „sehr zufrieden“ (5) einordnen. Bei „hoch“ (4) oder „sehr hoch“ (5) wird der Nutzer in der Auswertung als zufrieden eingestuft, andernfalls nicht. Wie viele Bilder der Nutzer wie lange in einem Task betrachtet hat, ist den Eye-Tracking-Videos zu entnehmen.

3 Literaturbericht

Untersuchungen des Suchverhaltens hinsichtlich des Ziels des Nutzers und dessen Experience (z. B. zufrieden oder unzufrieden) mit Hilfe von Interaktionsdaten liegen in verschiedenen Kontexten vor (z. B. Agichtein, 2012; Adafre, 2006). Agichtein et al. (2006) zeigen auch, dass das Suchziel Einfluss auf die inhaltlichen Präferenzen der Nutzer hat, allerdings liefern diese keinerlei Information darüber, welche Rolle die einzelnen Inhaltselemente dabei spielen. Die hier vorgestellte Studie baut außerdem auf bekannten Zusammenhängen zwischen Maus und Blickbewegung auf, die bei Rodden (2008) und Guo & Agichtein (2010-2) zu finden sind. Lerntheorien erklären die Effektivität einer multimedialen Gestaltung von Wikipedia-Artikeln mit Hilfe unterschiedlicher Formen der kognitiven Verarbeitung, basierend auf der Doppelkodierungstheorie von Paivio (1971). Die kognitive Theorie multimedialen Lernens (ursprünglich als SOI-Modell von Mayer (2001)) nimmt an, dass verbale und visuelle Informationen in verschiedenen Kanälen des Gedächtnisses verarbeitet werden. Durch die doppelte Verarbeitung der Informationen wirkt sich eine Kombination aus textlicher und bildlicher Informationspräsentation gegenüber einer rein textlichen als lernförderlich aus. Das Hinzufügen von Bildern zu Texten zeigt allerdings nur dann die angestrebte Wirkung, wenn die Bilder Funktionen erfüllen, wie z. B. eine darstellende, Organisations-, interpretative oder transformierende Funktion (Levin, 1981). Welche Text-Bild-Kombinationen für den Nutzer hilfreich sind, hängt vom jeweiligen Nutzungskontext ab. Verschafft der Leser sich einen kurzen Überblick über einen Artikel, ist eine Abbildung mit Organisationsfunktion hilfreicher als ein rein dekoratives Bild, das in diesem Kontext eher ablenkt. In Situationen, in denen der Nutzer Zeit hat (z. B. beim sog. browsing), können dekorative Bilder von Nutzen sein, zur attraktiveren Gestaltung des Artikels und zur größeren Nutzerzufriedenheit beitragen. Die Appraisal-Theorie (Arnold, 1960) besagt, dass Emotionen durch Bilder erst durch ihre persönliche Bedeutung für das Individuum ausgelöst werden. Sowohl der Nutzer, als auch der Task-Typ sind also in diesem Informationsprozess für die emotionale Aktion bzw. Reaktion auf Bilder ausschlaggebend.

4 Studie

Die Studie fand von April bis Juni 2012 an der Universität Regensburg statt.

In einer Laborstudie wurden Daten von 28 Testpersonen erhoben. Es handelte sich dabei in der Regel um

<i>Lookup</i>	Sie sehen zum hundertsten Jahrestag des Titanic Unglücks eine Reportage. Als Sie nach einer Werbepause zu spät zurückschalten, entgeht Ihnen die Information über die genaue Anzahl der Passagiere. Sie beschließen diese Information in der Wikipedia zu recherchieren.
<i>Learn</i>	Du hast heute Abend ein Date mit einer/m Lacrosse SpielerIn. Ihr wollt euch gemeinsam in einer Sportbar ein Spiel anschauen. Um genügend Gesprächsstoff für den Abend zu haben und weil du nicht als unwissend dastehen möchtest, informierst du dich über die grundlegenden Spielregeln und die wichtigsten Informationen zum Spiel. Die Wikipedia hat sich hier schon öfter als Nachschlagewerk profiliert.
<i>Casual-Leisure</i>	Man teilt dem Probanden mit, er befindet sich jetzt an einem Rechner im CIP Pool und hat bis seine Veranstaltung beginnt, noch ein paar Minuten Zeit, die er mit dem freien Surfen auf Wikipedia verbringen kann.

Tabelle 1: Beispielaufgaben.

action	label	description	
Read	RE	User is reading text	
Scan	SC	User scans content e.g. headlines, lists or whole page	
Examine	EX	User examines element	
Navigate	NV	User navigates	

element	label	element	label
Headline	HD	Text passage	TX
List	LI	Introduction	IN
Picture	PI	Info Box	IB
Charts, tables etc.	IG	Links in Wikipedia	WI
Other navigation	ON		

Abb. 1: Bedeutung der Videolabels.

erfahrene Wikipedia-Nutzer, die sich mit den Suchmöglichkeiten innerhalb der Wikipedia gut auskannten. Die Testpersonen waren beiderlei Geschlechts. Bisher wurden noch nicht alle Daten ausgewertet, doch die Ergebnisse reichen aus, um die Methodik zu evaluieren, zu verbessern und die Durchführbarkeit der Ideen zu testen und zu bestätigen.

Jeder Proband bearbeitete sechs Suchaufgaben, zu deren Lösung ausschließlich die Wikipedia verwendet werden durfte. Jede Aufgabe beinhaltete jeweils zwei Tasks der Typen *Lookup*, *Learn* und *Casual-Leisure* (Machionini, 2006). Sie wurden randomisiert zugeteilt, um einen möglichen Lerneffekt zu minimieren (Borlund, 2003). Die Aufgaben waren in einen Kontext eingebettet, enthielten also nicht nur eine Fragestellung, sondern zu-

sätzlich Informationen zum Zweck der Aufgabe und die fiktive Vorgeschichte dazu (siehe Tabelle 1 Beispielaufgaben). So wurde ein möglichst realistisches Szenario geschaffen, um bei allen Versuchspersonen eine ähnliche Vorstellung von der Aufgabe zu erzeugen. Damit die Versuchsperson nicht durch den Versuchsleiter beeinflusst wird, wurden die jeweils notwendigen Informationen während der Durchführung des Experiments auf dem Bildschirm der Testperson angezeigt.

5 Daten

Im Rahmen des Experiments wurden folgende Daten erhoben:

- Fragebögen vor und nach jedem Task sowie vor und nach der Studie, um demographische Angaben sowie die Zufriedenheit der Testperson mit der Durchführung der Aufgaben zu evaluieren.
- Eye-Tracking-Daten: Mit Hilfe eines SMI RED Eye-Trackers wurde die Blickinteraktion der Testperson mit der Seite als Video und in Form von Log-Files, die Angaben zu Fixationen und Sakkaden enthalten, aufgezeichnet. Die Videos wurden manuell vollständig mit zweierlei Labels annotiert (s. Abb. 1): Das erste Label beschreibt die Aktion, die die Testperson gerade durchführt (z. B. Lesen oder Betrachten), das zweite das Inhaltselement, welches gerade im Fokus ist (z. B. ein Bild oder ein Artikelabsatz).
- Firefox-Logs: Es werden alle Interaktionen der Testperson mit dem Browser incl. der jeweiligen Metadaten dazu wie z. B. die *current position* und der *time stamp* während des Suchprozesses aufgezeichnet (events wie mousemove, keypress, scroll etc.)
- Elektromyographische (EMG) Daten: Mit einem EMG-Gerät ist die Messung von Muskelaktivität durch die Aufzeichnung elektrischer Spannung möglich. Während der hier beschriebenen Experimente wurden Werte für die Aktivität zweier Gesichtsmuskeln aufgezeichnet, den corrugator und den zygomaticus maior, die Indikatoren für emotionale Reaktionen sind (Brown, 1980).

Um eine Synchronisation aller Daten zu ermöglichen wurden jeweils die Zeiten in ms mitprotokolliert.

6 Auswertung

Die Auswertung der Daten erfolgte hinsichtlich verschiedener Fragestellungen.

6.1 Vorhersagbarkeit des Task-Typs

Die Analyse des Nutzerverhaltens ergab Hinweise auf signifikante Unterschiede bei der Verwendung von Inhaltselementen während der unterschiedlichen Task-Typen, wie es im paarweisen Vergleich in Abbildung 2 deutlich wird. Werden bei *Look up* eher Tabellen und grafische Darstellungen genutzt, konzentriert sich der Nutzer bei *Learn*-Aufgaben mehr auf Einleitung und Listen. Im Bereich *Casual-Leisure* wurden sehr viele verschiedene Inhaltselemente verwendet.

Zudem führen verschiedene Aufgaben auch zu unterschiedlichen Mustern im Leseverhalten. Bei *Look up*-Aufgaben wird der Text nur kurz überflogen, dagegen werden bei den anderen Aufgabentypen viele Textpassagen gelesen. Abbildung 2 zeigt eine signifikant unterschiedliche Verteilung der annotierten Video-Labels bei den verschiedenen Aufgabentypen.

action	LO vs RE		LO vs CA		RE vs CA	
	χ^2	p-value	χ^2	p-value	χ^2	p-value
EX	9	0.011	9	0.029	18	0.006
NV	9	0.011	9	0.011	18	0.001
RE	13	0.043	6	0.301	27	0.079
SC	36.563	0.064	45	0.039	45	0.039

Abb. 2: Chi-Quadrat-Test für die Verteilung der Labels für die verschiedenen Task-Typen (LO: Lookup, RE: Learn, CA: casual-leisure).

6.2 Korrelation der Browser-Interaktion mit den Videolabels

Die Eye-Tracker-Videos wurden per Hand annotiert. Somit kann eine Analyse der Nutzeraktionen immer erst nach Abschluss eines Tasks stattfinden, außerdem ist der Einsatz eines Eye-Trackers in der Regel nur in Laborstudien möglich. Deshalb ist es ein Ziel, allein anhand der Browser-Interaktion die Nutzer-Aktionen zu erkennen. Dazu wurde die Korrelation der Video-Labels mit den Firefox-Logdaten überprüft. Zunächst wurden die einfacheren Events wie *mousemove*, *click* und die Blickbewegungen des Nutzers auf signifikante Korrelationen über-

prüft. Die Auswertung der ersten Daten zeigt, dass es möglich ist, bestimmte Nutzer-Aktionen (entsprechend den Video-Labels) allein aus der Browser-Interaktion zu erkennen. Da Browser-Logs einfach auszuwerten sind, wäre auch eine Berechnung zur Laufzeit vorstellbar.

6.3 Zufriedenheit in Abhängigkeit von Inhaltselementen

Die Auswertung hinsichtlich der Zufriedenheit der Nutzer wurde zunächst auf die Untersuchung des Einflusses von Bildern beschränkt. Bei den Aufgabentypen *Learn* und *Casual-Leisure* werden Bilder sehr viel häufiger verwendet. Hier geht es um die Aneignung von Wissen, nicht aber um das Nachschlagen ganz konkreter Informationsstücke (siehe *Look up*). Bilder stellen dabei eine zusätzliche Informationsquelle oder eine Visualisierung des Textes dar (Bsp. Spielfeld des Spiels Lacrosse). Beim Task-Typ *Look up* werden Bilder fast nie betrachtet. Da die Aufgabenstellung lediglich auf Textinformation abzielt, ist es erklärlich, dass der Nutzer die Bildelemente nicht als erste Informationsquelle heranzieht. Die Fragestellung bei *Look up 1* nach der Anzahl der beim Untergang der Titanic an Bord befindlichen Passagiere suggeriert den betreffenden Artikelabschnitt bzw. die entsprechende Infobox zu lesen. Es ist anzunehmen, dass sich bei veränderter Fragestellung, die zum Beispiel auf Informationen abzielt, die eher in einem Bild zu finden sind, auch beim Aufgabentyp *Look up* eine stärkere Bildverwendung nachweisen lassen würde. Allgemein hängt die Bildverwendung vermutlich neben dem Fachgebiet und dem Task-Typ auch von der konkreten Fragestellung ab.

Basierend auf der Angabe zu ihrer subjektiven Zufriedenheit nach jeder einzelnen Aufgabe im Post-Task-Fragebogen waren die Nutzer bei allen Aufgabentypen unabhängig von der Bildanzahl gleich zufrieden oder unzufrieden mit ihrem Suchverlauf. Bei *Casual-Leisure* konnte dieses Merkmal nicht erhoben werden, weil dieser Task-Typ keine explizite Fragestellung beinhaltet. Um weitere Erkenntnisse über die Zufriedenheit des Nutzers in Abhängigkeit von der Bildverwendung zu gewinnen, werden im

action	mousemove		element	mousemove		Task	scroll		click		mousemove		avg.sacc.dist	
	high	low		high	low		action	el.	act.	el.	act.	el.	act.	el.
NV	5	6	IN	30	12	1	***	***	***	***	***	***	***	**
RE	18	5	IB	8	10	2	***	***				*	***	***
SC	41	18	WI	5	6	3	*	**		*	***	***	*	**
			LI	21	1	4	*	***	*			**		***
						5	***	***			***	***	*	
						6	***	***	**		***	***	**	***

Abb. 3: Gleichzeitiges Auftreten bestimmter Nutzeraktionen mit *mousemove*, rechts die Ergebnisse des Chi-Quadrat-Tests für die Korrelation der einfachen Browser-Events mit den Videolabels.

weiteren Verlauf der Datenanalyse die mit Hilfe des EMG-Gerätes erfassten emotionalen Reaktionen herangezogen.

7 Ausblick

Die bisher durchgeführte Analyse der Daten zeigt bereits sehr interessante Ergebnisse. Es stellte sich heraus, dass vermutete Zusammenhänge nachweisbar sind, wie etwa die unterschiedlichen Präferenzen für Inhaltselemente bei verschiedenen Suchaufgaben. Es ist vorstellbar, weitere detaillierte Untersuchungen in dieser Hinsicht vorzunehmen, zum Beispiel ob eine Abhängigkeit der Präferenzen vom Fachgebiet oder dem aktuellen emotionalen Zustand des Nutzers zu erkennen ist. Vermutlich beeinflussen noch weitere Faktoren, die in diesem Experiment nicht berücksichtigt wurden, die Inhaltspräferenzen der Nutzer (Vorkenntnis, Geschlecht, Alter, Erfahrung, benutztes Endgerät, grundsätzliche Präferenzen, Einflüsse durch Sehschwächen, etc.). Vor allem der Einsatz und der Gebrauch der Bilder bzw. der unterschiedlichen Bildtypen ist ein großes Forschungsfeld.

Literatur

- Adafre, S. F. & de Rijke, M. (2006). Exploratory Search in Wikipedia. In: Proceedings of Special Interest Group on Information Retrieval 2006, Workshop on Evaluating Exploratory Search Systems.
- Agichtein, E. & Brill, S. D. R. (2012). Mouse Tracking: Measuring and Predicting User's Experience of Web-based Content. In: Proceedings ACM Conference on Human Factors in Computing Systems 2012.
- Agichtein, E.; Brill, E.; Dumais, S. & Ragno, R. (2006). Learning user interaction models for predicting web search result preferences. In: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval. New York, NY, USA: ACM, pp. 3–10.
- Arnold, M. (1960). Emotion and Personality. 1. Bd.: Psychological Aspects. New York.
- Borlund, P. (2003). The IR evaluation model: a framework for evaluation of interactive information retrieval systems. In: Information Research 8(3), pp. 1–16.
- Buscher, G.; Dengel, A. & Van Elst, L. (2008). Eye movements as implicit relevance feedback. In: CHI'08: Extended Abstracts on Human Factors in Computing Systems, pp. 2991–2996.
- Claypool, M.; Le, P.; Wased, M. & Brown, D. (2001). Implicit interest indicators. In: Proceedings of the IUI, pp. 36–40.
- Guo, Q. & Agichtein, E. (2010). Ready to buy or just browsing?: detecting web searcher goals from interaction data. In: Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. New York, NY, USA: ACM, pp. 130–137.
- Levin, J. R. (1981). On functions of pictures in prose. In: F. J. Pirozolo & M. C. Wittrock (Eds.): Neuropsychological and cognitive processes in reading, pp. 202–228. New York: Academic Press.
- Kappas, A. & Müller, M. G. (2006). Bild und Emotion – ein neues Forschungsfeld. Theoretische Ansätze aus Emotionspsychologie, Bildwissenschaft und visueller Kommunikationsforschung. In: Publizistik 51(1), S. 3–23.
- Marchionini, G. (2006). Exploratory search: from finding to understanding. In: Communications of the ACM 49(4), pp. 41–46.
- Mayer, R. E. (2001). Multimedia Learning. New York: Cambridge University Press.
- Paivio, A. (1971). Imagery and Verbal Processes. New York: Holt, Rinehart and Winston.
- Rodden, K.; Fu, X.; Aula, A. & Spiro, I. (2008). Eye-mouse coordination patterns on web search results pages. In: CHI '08 extended abstracts on Human factors in computing systems. New York, NY, USA: ACM, pp. 2997–3002.
- White, R. W. & Kelly, D. (2006). A study on the effects of personalization and task information on implicit feedback performance. In: Proceedings of CIKM 2006, pp. 297–306.



Hanna Knäusl, M. A.
Lehrstuhl für Informationswissenschaft
Universität Regensburg
Universitätsstraße 31
93053 Regensburg
hanna.knaeusl@ur.de

studierte von 2002 bis 2008 Informationswissenschaft und Politikwissenschaft an der Universität Regensburg. Seit 2009 ist sie Wissenschaftliche Mitarbeiterin am Lehrstuhl für Informationswissenschaft. Ihre Forschungsschwerpunkte liegen im Bereich der Wikipedia-Forschung und dem Information Retrieval. Sie ist seit 2004 Mitglied im Vorstand des Hochschulverbands Informationswissenschaft (HI e. V.), seit 2011 stellvertretende Vorsitzende.



Barbara Rösch, M. A.
Lehrstuhl für Informationswissenschaft
Universität Regensburg
Universitätsstraße 31
93053 Regensburg
barbara.roesch@stud.uni-regensburg.de

hat Pädagogik und Informationswissenschaft studiert und arbeitet seit April 2012 als WHK am Lehrstuhl für Informationswissenschaft an der Universität Regensburg. Dort betreut sie das Eye-Tracking-Labor und betreibt im Zuge ihrer Promotion Forschung zum Thema „Text-Bild-Verhältnis in Wikipedia“.



Lea Schubart, B.A.
Lehrstuhl für Informationswissenschaft
Universität Regensburg
Universitätsstraße 31
93053 Regensburg
lea.schubart@stud.uni-regensburg.de

ist wissenschaftliche Hilfskraft am Lehrstuhl für Informationswissenschaft.